DATA VALIDATION FOR DATA BASES OF
THE NATIONAL PERFORMANCE AUDIT PROGRAM

by

John Holland

ManTech Environmental Technology, Inc.
Research Triangle Park, NC

CAUTION

Disclaimer: This Standard Operating Procedure has been developed for use by ManTech Environmental Technology, Inc. in support of the National Performance Audit Program (NPAP) under contract to the U.S. Environmental Protection Agency and may not be applicable to the activities of other organizations.

Approved by:

_____

Kenneth J. Caviston, Manager   Date
ManTech Environmental Technology, Inc.

_____

Mark Shanis   Date
EPA work Assignment Manager

Effective: When approved

PROCEDURAL SECTION

1.0   **SCOPE AND APPLICATION**

1.1   This procedure is intended for use in performing data validation on National Performance Audit Program (NPAP) data bases.

1.2   It is performed before doing a data summary of a quarter's audit data or an annual report.

1.3   It is not intended to be a statistical validation, but a method of identifying and checking all audit data that result in a large percentage error.

2.0   **SUMMARY AND METHOD**

2.1   For all audits except VOCs and carbonyls, a data summary is calculated for each audit and quarter to be validated.

2.2   The data summary consists of mean percent differences and standard deviations for each audit concentration level in each U.S. EPA Region.

2.3   If any mean or standard deviation is 20% or more, all data sets included in that mean or standard deviation are checked.

2.4   For VOCs and carbonyls all data sets are checked by comparing the original data sheets to the values stored in the NPAP data system.

2.5   AIRS site codes are validated prior to the annual AIRS site report by running a computer program that lists bad site codes and checking that list.

3.0   **PROCEDURE for Lead, Hi-Vol/SSI, Dichot, Ozone, Nitric Oxide, Carbon Monoxide, Sulfur Dioxide, and Nitrogen Dioxide**

3.1   Select the calendar year of the audit to be validated.
Enter your password.
❗   Select "4. Print Lists."
❗   Select "3. Lists Reporting Results."
❗   Select output preference from the list (usually to the printer).
❗   Select the audit from the next list on the screen.
❗   Enter the calendar quarter.
❗   Select menu selection "3. List Means and Standard Deviations by Quarter."
❗   Answer "N" (no) when the system asks "Remove outliers?".
❗   The data will be accumulated and the summary report will be output.
❗   Type "0" (the number zero) to exit.

3.2   If any means or standard deviations listed are over 20%,
- Check all data for that level, quarter, and U.S. EPA Region to verify proper entry from the original data forms.
- Make written notations in the data validation notebook or on the report by the mean or standard deviation of each group of data checked.
- If notes are written on the report, place it in the data validation file or attach it to a page in the notebook.

3.3   Correct any errors by following *NPAP-SOP-007, Editing National Performance Audit Program Data Bases*.

3.4   Record the date the validation was completed and who did the validation in the data validation notebook and in the computer data system. The data validation software is accessed from the **Main Menu**:
- Select "7. Mark Data Validation."
- Select "3. Mark quarters validated."
- Select the audit from list on the screen.
- Enter the calendar quarter.
- Enter who performed the data validation and the date the data validation was completed. The system will return to the audit selection menu.
- Type "1" (the number one) to view data validation status for all audits, or type "0" (the number zero) to exit.

4.0   <u>**PROCEDURE for VOCs and Carbonyls Audits**</u>

4.1   Perform the data validation before any data are sent to participants.

4.2   Check all data entered for the audit by comparing the computer report to the original data sheet.

4.3   Correct any errors by following *NPAP-SOP-007, Editing National Performance Audit Program Data Bases*.

4.4   Make written notations in the data validation notebook of the date, who did the data validation and what the findings were (i.e. "all data were correct", "one AIRS site number was in error", *etc*.).

4.5   Record the date the validation was completed and who did the validation in the data validation notebook and in the computer data system. The data validation software is accessed from the **Main Menu**:
- Select "7. Mark Data Validation"
- Select "3. Mark quarters validated"
- Select the audit from list on the screen.

! Enter the calendar quarter.
! Enter who performed the data validation and the date the data validation was completed. The system will return to the audit selection menu.
! Type "1" (the number one) to view data validation status for all audits, or type "0" (the number zero) to exit.

5.0   PROCEDURE FOR AIRS SITE CODES

5.1   Select "7. Data Validation" on the Main Menu.

5.2   Run the AIRS site data validation program for last year and the current year.

5.3   Send the outputs to convenient local files. The program creates lists of possibly incorrect AIRS site codes from all the audit data records.

5.4   Edit the files to reduce the number of pages and print them.

5.5   Examine the lists for possible data entry errors.

5.6   Check all possible errors with the original data sheets.

5.7   Correct all errors in the computer Data System following NPAP-SOP-007, Editing NPAP Data Bases.

6.0   QUALITY CONTROL, AND QUALITY ASSURANCE SECTION

6.1   This SOP is part of the internal QA/QC on NPAP data management.

6.2   The activities specified in this SOP are subject to EPA systems Audits.

REFERENCES

1.   NPAP-SOP-007, Editing National Performance Audit Program Data Bases.

2.   NPAP-SOP-007, Editing National Performance Audit Program Data Bases.